


**Corporate**

Foundation  
Beliefs  
Management Team  
Advisory Board  
Board Members  
Partner With Us

**Press**
**Investor**
**Contact Us**
**Search**

☐ Site ☒ School


**Find a Course/Tutor**

SAT

Enter Zip/Postal Code


**Additional Options**

800-2REVIEW

## Testing the Testers 2003: An Annual Ranking of State Accountability Systems Executive Summary

During the Winter of 2002-2003, The Princeton Review conducted Testing the Testers 2003, its second Annual Ranking of State Accountability Systems. Unlike other studies, ours is not primarily concerned with the rigor of academic standards or of the tests that measure them. Rather we focused on the policies that determine the overall character and effectiveness of each accountability system. Properly conceived and well-implemented, these policies will tend to produce systems that are consistent, secure, open to public scrutiny, and flexible enough to improve over time. We also believe they will tend to encourage and support an evolution to better and more effective schools.

[View State by State Rankings](#)

[Download the Complete Study, Data Tables, or Spreadsheets](#)

**Get more information:**

[About Us](#)  
[Core Beliefs](#)  
[K-12 Services](#)

As the stakes for testing rise, and with the pressure of the Federal No Child Left Behind act (NCLB), accountability systems increasingly affect what gets taught and how. As a result they will strongly influence how schools develop over the next several years. Simply put, good accountability systems will tend to result in better schools, and bad systems will create worse ones. The purpose of Testing the Testers is to highlight good and bad accountability practice with the hope of helping the overall tide to rise. By "good" we mean accountability systems that will lead not only to improvement on test scores as well as on other measures of school quality, that will support educator professionalization, make school a more satisfying and rewarding experience for students, and importantly, that will be able to improve and adapt as political and pedagogical realities change. Raising test scores is not that difficult if raising scores is all you want to do, and are willing to sacrifice the rest of what school means in order to do so. That, to us, would be bad accountability.

We collected data on twenty-two relevant indicators from each state and the District of Columbia. Each indicator was grouped in one of four major criteria and states received a score of either zero, one, or two points depending upon how their program performed. The criteria were:

1. **Academic Alignment:** High-stakes tests are aligned to academic content knowledge and skills as specified by the states' curriculum standards.
2. **Test Quality:** The tests are capable of determining that those curriculum standards have been met.
3. **Sunshine:** The policies and procedures surrounding the tests are open, and open to ongoing improvement.
4. **Policy:** Accountability systems will tend to affect education in a way that is consistent with the goals of the state.

These criteria were weighted at 20%, 15%, 30%, and 35% respectively and the raw scores scaled accordingly to give each state and the District of Columbia a ranking from one to fifty-one (the highest possible weighted score was 100). Each state was also assigned letter grades on the A-F scale for each of the four criteria.

**The best programs are:**

Rank	State	Weighted Score	Alignment 20%	Test Quality 15%	Sunshine 30%	Policy 35%
1	NY	88.5	B+	A	B	A-
2	MA	85.7	B-	A	A-	B+
3	TX	84.3	B-	B+	A-	A-
4	NC	84.0	B-	A	B	A-
5	VA	81.7	A	A	B+	B-
6	LA	81.0	B-	A	B+	B+
6	FL	81.0	B-	A	B+	B+
8	AZ	80.2	B-	A	C+	A-
8	OK	80.2	B-	A	B	B+
10	CA	79.7	B+	A	B	B-

**The worst programs are:**

Rank	State	Weighted Score	Alignment 20%	Test Quality 15%	Sunshine 30%	Policy 35%
41	KS	58.2	D	A	C+	C+
42	IN	56.8	D	A	C	C+
43	HI	55.5	C-	B+	C-	B-
44	WY	54.5	F	A	C	B-
45	ND	54.3	C-	B+	C-	C+
46	WI	53.2	C-	A	C-	C+
47	WV	52.2	D	A-	F	B-
48	SD	49.8	B-	A	F	C
49	RI	48.5	C-	A	F	B-
50	MT	29.0	F	B-	F	C-

Only Virginia received two A's, and no state received an A for either of our most significant criteria, Sunshine and Policy. Nearly 30% of states received overall scores of 65 or lower, and of the individual grades given to the bottom-performing twenty states, nearly 40% were C or lower. On the positive side, forty-six programs received grades of B+ or better for the quality of the test instruments themselves, with only Utah scoring lower than a B-.

Although the rankings are affected by the weighting we applied (especially for those states in the middle three quintiles) most states tend to do things well or poorly with some consistency across all indicators, regardless of weighting. Most reasonable weightings (including no weighting at all) do not drastically alter the composition of the top or bottom rankings. Rankings for unscaled scores are presented in the body of this report, and readers are encouraged to download the data spreadsheet from [here](#) and formulate their own weightings and judgments.

For More Information, contact: Harriet Brand, (212) 874-8282 ext. 1091, [harrietb@review.com](mailto:harrietb@review.com) or Robin Raskin, (212) 874-8282, ext. 1649, [robinr@review.com](mailto:robinr@review.com)

[Download the Complete Study, Data Tables, or Spreadsheets](#)

**Get more information:**

[About Us](#)

[K-12 Services](#)

[Privacy Policy](#) | [Terms & Conditions](#) | [Site Map](#) | [Employment](#) | [Company Information](#) | [Contact Us](#)

Copyright Notice

SAT | PSAT | ACT | GMAT | GRE | LSAT | MCAT | USMLE

